DOCUMENT RESUME

ED 269 448 TM 860 282

AUTHOR Winters, Lynn
TITLE Technical Trouble Spots in the California High School Quality Indicator Program.
INSTITUTION Palos Verdes Peninsula Unified School District, CA.
PUB DATE [May 85]
NOTE 16p.; Paper presented at the Los Angeles County School Board Conference on Quality Indicators (Los Angeles, May, 1985).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accountability; Achievement Gains; *Educational Quality; High Schools; Instructional Improvement; Measurement Techniques; *Research Problems; Research Projects; Research Utilization; *School Effectiveness; Scoring; State Legislation; State Standards; Testing Problems; *Validity
IDENTIFIERS California High Schools Accountability Program; *California High Schools Performance Reports; *California High Schools Quality Indicator Program

ABSTRACT

The purpose of the California High School Performance Reports is to provide a means for the public to judge a school's progress towards: (1) implementing the educational reforms outlined in California State Legislature Senate Bill 813; (2) upgrading the quality of its instructional program; and (3) improving student achievement. This paper discusses the technical inadequacy of the reports. The most serious defects of the performance reports are that they do not measure what they purport to measure and the data provide no information for judging school quality. The quality indicators are uninterpretable, can be misleading, and are useless for determining how well schools are meeting the mandates of the reform legislation. Although the reports offer two standards against which the public is supposed to measure school quality (the state targets and comparison band percentile ranks) these standards have no relationship to the stated aims of the High School Accountability Program (HSAP). A review of these two indicators leads to the conclusion that each is an inappropriate measure of schooling. The need for HSAP to publish data related to worthwhile educational goals based upon a consensual definition of school excellence is emphasized. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED269448

Technical Trouble Spots in the California High School Quality Indicator Program

Lynn Winters
Director of Program Evaluation and Research
Palos Verdes Peninsula Unified School District

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L. Winters

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM 860 282

Technical Trouble Spots in the California High School Quality Indicator Program

Lynn Winters
Director of Program Evaluation and Research
Palos Verdes Peninsula Unified School District

The technical inadequacy of the California High School Performance Reports is well known. Such problems as the inaccuracy of CBEDS and CAP questionnaire data, the inappropriateness of using SAT scores, the differential bases for calculating course enrollments, and the reporting of unweighted college grade point averages represent the problems most often cited in critiques of the quality indicators. These technical shortcomings represent the kinds of problems most easily remedied and least related to the two most serious defects of the performance reports: They don't measure what they purport to measure and the data provide no information for judging school quality.

Lack of Validity

The stated purposes of the performance reports are to provide a means for the public to judge a school's progress towards 1) implementing the educational reforms outlined in SB 813, 2) upgrading the quality of its instructional program and 3) improving student achievement. Unfortunately, the data reported bear no relationship to these three purposes.

In the best of situations, the quality indicators are uninterpretable; in the worst they are totally misleading. In

either case, they are useless for determining how well schools are meeting the mandates of the reform legislation.

School reforms proposed in SB 813 specify changes in the procedures and inputs to schooling, not the outcomes. The High School Performance Reports, on the other hand, report on the outcomes of schooling. (See Table 1). These outcomes, however, are tied neither to the provisions of SB 813 nor to some other statement of desired instructional goals. If the reports were designed to measure how well schools are implementing the SB 813 reforms, they would present data related to 813 mandates. Instead, the Performance Reports publish California Assessment Program (CAP) results, SAT averages, and percentages of students enrolled in specified courses (none of which correspond to the pattern of courses mandated by SB 813 appearing in Table 2).

Given that the performance reports don't address how well schools are implementing SB 813 reforms, do the reports meet their other stated aims: published to describe the quality of instruction and student achievement at a particular school? The answer is "no". Although the reports offer two standards against which the public is supposed to measure school quality, the state targets and comparison band percentile ranks, these standards, in fact, have no relationship to the stated aims of the high school accountability program. The next section reviews these two indicators in depth and posits that each is an inappropriate measure of schooling.

Table 1

Summary of 1983 High School
Accountability Program Quality Indicators

Quality Indicator	Statewide Targets 85-86	87-88	89-90
A. Course Enrollments			
1. Mathematics - 3 or more years			
2. English - 4 or more years			
3. Science - 3 or more years			
4. History/Social Science - 4 or more years			
5. Foreign Language - 3 or more years			
6. Fine Arts - 1 or more years			
B. California Assessment Program (CAP) Scores			
1. Reading			
Percent correct score	62.7	63.7	64.7
Percent scoring			
Above Q	52*	56*	60*
Above Q	26*	27*	28*
Below Q	24*	22*	20*
2. Mathematics			
Percent correct score	67.9	68.9	69.8
Percent scoring			
Above Q	52*	56*	60*
Above Q	26*	27*	28*
Below Q	24*	22*	20*
3. Writing			
% of students completing one writing assignment per week (data to be collected and reported by individual schools)	50	75	100
4. Homework			
% of students completing:			
- one hour of homework per night	66	80	90
- two hours of homework per night. (data to be collected and reported by individual schools)	30	40	50

*Targets are compared to 1983-84 standards.

Table 2

Senate Bill 813, Chapter 498 of the Statutes of 1983: The Hughes-Hart Education Act of 1983. Sacramento: California State Legislature, 1983, 214pp.

SEC.2. The Legislature believes that our schools should:

(a) Provide a variety of instructional styles and classroom settings which accommodate the differing ways that research has shown children learn.

(b) Maintain orderly and efficient school campuses which encourage positive attitudes among students and high morale and high quality teaching from teachers.

(c) Ensure that the specialized needs of identified groups of students are met by the effective use of categorical aid funding.

(d) Provide appropriate and meaningful instruction to meet the variety of future and career goals of students.

(e) Assure that pupils achieve academic proficiency in the essential areas of skill and knowledge.

(f) Develop each pupil's sense of respect of self and others, personal and social responsibility, and critical thinking.

(g) Involve parents and community members in a board range of activities at each school, recognizing the vital role parental attitudes and values have in their children's education.

Curriculum, Instruction, Standards

Commencing with the 1986-87 school year, no pupil will receive a high school diploma who has not in grades 9-12 completed:

- 3 years of English
- 2 years of mathematics
- 2 years of science (biological and physical)
- 3 years of social studies (U.S. History and geography; world history, culture, and geography; American government, civics, and economics)
- 1 year of fine arts or foreign language
- 2 years of physical education (unless exempted)
- Such other coursework as the district may specify (Sec. 94)

State Targets

The purpose of publishing state targets for CAP scores, course enrollments, and SAT scores is so that school districts will know where they stand with reference to a desired goal. Unfortunately, these desired goals (state targets) are not based on any clear definition of school excellence. In fact, schools can meet or exceed the targets without making real improvements in their instructional program. This flaw in the data is best exemplified by state CAP score goals.

Schools are expected to raise their CAP scores by 1% a year. This "gain", however, is less than the standard error of measurement for schools scoring at the extreme ends of the state score distribution. Such gains could be observed by chance alone without "real" changes having occurred in the instructional program. Remarkable increases in grade 12 CAP scores between December, 1983, and December, 1984, illustrate the fallacy of equating score gains to an improved instructional program. In schools where dramatic gains occurred, increases were due to improved testing conditions, incentives, and better test preparation. It is hardly possible for program changes made in the last six months of a student's junior year and the first four months of the senior year to have a dramatic effect on developmental skills, such as reading, which take years to acquire. It is even less probable that schools actually revamped their curriculum mid-year, 1983, in order to remediate the basic skills of the Class of 1985. CAP results weren't even published until May, 1984. Also, most schools spent the summer of 1984,

developing SB 813 course expectancies. Development of expectancies would precede any program changes for students taking the 1984 CAP. The earliest curricular change that could have occurred was fall, 1984. Thus, CAP test gains, at best, measure four months of an "improved curriculum," hardly enough to account for increases of 2% or more in test sccres.

A second state target focuses on CAP score distributions. Yet score distributions reflect student background characteristics as well as levels of achievement. How much of a change in the score distribution from year to year is attributable to improved instruction and how much is attributable to population differences? What happens when test content is changed, as is contemplated for the 1986 CAP test? Shifts in score distribution are uninterpretable when test content or student population varies, yet these changes are offered as evidence of improved instructional quality.

The state targets CAP score distributions specify the percentage of students who should score in selected quartile ranges on the reading and math sections of the CAP. The state expects a 1% a year increase in the number of students scoring at/above the 75th percentile, a 4% increase in the number scoring at/above the 50th percentile and a 2% decrease in the number scoring at/below the 25th percentile. How were these targets set? Why is it more reasonable to expect a 2% decline in the lowest quartile and only a 1% increase in the top quartile? Without a rationale and evidence that targets represent real increases in student achievement, these data are useless for assessing school quality.

State targets have also been set for course enrollments, hours of homework and number of writing assignments. The same question applies: what is the basis for the standard? All standards are arbitrary, but reasonable standards are based on informed judgment. What research or practice indicates that the number of hours of homework ought to be the same for third graders as for twelfth graders? What is the positive effect on student achievement of completing one writing assignment per week when writing topics, essay length, and criteria for judging the assignment aren't specified? The quality indicators are operational definitions of some goal, but what is the goal? How are state targets for course enrollments, SAT scores, CAP scores, and the other quality indicators related to improved instruction and to better preparation for employment or college?

Comparison Bands

If the state targets have little meaning for assessing program worth, comparison band percentile ranks have none at all. In the past, CAP scores have been reported to the public with percentile ranks comparing a given school to other schools in California. A school receiving a state percentile rank of 70, for example, scored higher than 69% of the schools in the state. The performance reports still assign percentile ranks to schools, but not according to relative position among all high schools statewide. Instead, the ranks compare "similar" schools, schools serving students whose parents have similar levels of education (as reported by the students).

The assumption underlying the comparison band is that parent education level (defined as the number of years of education of one parent) is related to school excellence. Thus, schools serving pupils whose parents have many years of education are more likely to score higher on state achievement tests and to offer a more challenging curriculum than schools serving parents with little formal education. This assumption simply does not hold.

Table 3 challenges the assertion that comparison bands group schools into similar levels of parent education. The five comparison groups were created by dividing the total number of high schools for which performance reports were generated by five, so as to make five groups roughly equal in size. The result of this procedure was to create groups equal in size but quite unequal in ranges of parent education. The difference in average parent education level in Band C is only 0.25, whereas the difference in Bands A and E is over 1.0. A more accurate procedure for sorting schools into similar categories would have been to group schools whose parent indices predicted similar CAP scores, course enrollments, and SAT scores. If bands had been created in this manner, there may have been 200 schools in one band and perhaps 20 in another, but the schools would have been "alike" in the way their demographic characteristics predicted student outcomes.

The existence of five different comparison groups, Bands A-E, also suggests that schools in one group differ significantly in

Table 3

Values of Parent Education Index of Comparison Bands:
Grade 12 CAP, December, 1983

Groups	Range of Parent Ed. Index		
Comparison Group	Low	High	Span
E	3.45	4.47	1.02
D	3.13	3.44	.31
C	2.87	3.12	.25
B	2.54	2.86	.32
A	1.10	2.53	1.43

Values
5 = advanced degree
4 = college graduate
3 = some college
2 = high school graduate
1 = not finish high school

Table 4

A Comparison of the Range of State Percentile Ranks in Each Comparison Band: CAP, December, 1983

	State Percentile Ranges: Low to High				
BAND	A	B	C	D	E
Reading	1-99	1-97	2-99*	11-99	13-99
Math	1-99*	1-85	1-97	5-99	4-99

*School in this group has highest score possible.

some way from schools in another group. Yet, because of the way the bands were created there is much overlap in performance among schools in different bands. From Table 4, it is apparent that for any one band, one can observe almost the entire range of CAP test performance. The table displays the lowest and highest state percentile rank received by schools in the band. With the exception of the schools in Band B, there were schools receiving a state rank of 99th percentile in all bands. All bands included schools that scored near the bottom of the state distribution.

While the existence of five different groups implies that qualitative differences exist among the comparison bands, it obscures the fact that schools in the "lower" Bands (A-B) may do as well or better than schools in "higher" Bands (D-E). At least one school in Band A had the highest possible CAP score in math; the same is true in reading for at least one school in Band C. Will the public, however, think that schools ranking at the 99th percentile in Band A perform as well as schools ranking at the 99th percentile in Band E?

Table 5 helps refute the assumption that significant qualitative differences necessarily exist between schools grouped in Band A, and Bands C, D, or E. This table displays the range of the percentage of students scoring at/below the 25th percentile on the Grade 12 CAP, December, 1983. Although higher bands have a smaller percentage of low scoring students, they don't win by much. The only schools having no scores below the 25th percentile appear in Band A. A hefty percentage of Band E

students (47%) scored below the 25th percentile in math. Yet the public may believe that Band A encompasses the "poor" and low scoring schools while E contains the high socio-economic, high scoring schools. In fact, the entire range of school "quality" (as defined by the state indicators) may occur in every band.

How necessary is it to create five comparison bands? Table 6 displays the state percentile rank equivalent to the median score (50th percentile) in each of the comparison bands. While it is true that the higher bands have higher median scores, it is also true (with three exceptions) that half of the schools in each band score at or above the state average. The discrepancy between the median reading and math scores in Bands A-B suggests that the low scores in those bands are related more to student population differences (i.e. bilingual students) than to the quality of the program offered. When half of the schools in four of the five comparison groups score above the state average in math, is it fair to say that there are five qualitatively different groups of schools in the state?

Table 6 also shows the invidiousness of ranking schools from 1-99 within comparison bands. In Band E, for example, 50% of the schools placed at/above a state rank of 98th percentile statewide in math. Yet these schools will be assigned comparison band percentile ranks ranging from 50-99 regardless of the fact they score in the upper 2% of state high schools.

Table 5

The Range of Percentage of Students Scoring at/below the 25th Percentile on the CAP, Grade 12, 1983, by Comparison Band

	Percent students in lowest quartile				
BAND	A*	B	C	D	E**
CAP Reading	0-100%	5-57%	4-49%	5-37%	4-34%
CAP Math	0-91%	6-66%	5-59%	1-44%	4-47%

*The only school having no one scoring at/below the 25th percentile was in Band A.

**Even in Band E, "the highest" band, at some schools almost as high a percentage of students score in the lowest quartile as in Bands C-D.

Table 6

State Percentile Rank Equivalent to the Score Representing the 50th Percentile in Each Comparison Band

BAND	A	B	C	D	E
Reading	14	31	50	63	86
Math	41	60	82	90	98

(With the exception of the schools in the boxed area, 50% of the schools in each band scored at/above the state average).

Table 7

Ranking of Bands from Highest to Lowest by the Percentage of Students Enrolled in Selected Courses

Rank	4 years social studies	4 years English	Advanced Math
1	B	E	E
2	A	B	A
3	C	D	C
4	D	C-A	D
5	E	---	B

Table 7 emphasizes the futility of using comparison bands to interpret information about schools. The table was constructed by ranking the bands from highest to lowest according to the value of selected indicators equal to the 90th percentile of each band. For example, a school enrolling 83.6% or more of its students in four or more years of social studies would stand at the 90th percentile in Band A. The 90th percentile for Band B social studies enrollment is 84.7%. The 90th percentile in Band E is 76.3% The table clearly shows that parent education level as defined by the comparison bands does not predict courses taken in high school. Bands with higher average levels of parent education don't necessarily have more students completing four years of English or social studies. In fact, the percent of students reporting they took four years of social studies is almost inversely related to level of education. If, then, there is little relationship between band membership and performance on the state-selected quality indicators, why separate schools into different groups at all?

Comparison bands have "halo" and "horn" effects. To be at the 99th percentile in Band A means to be the "best of the worst" in some minds. With the state not providing statewide ranks for the 1984-85 performance reports, the myth of the comparison band will be even more difficult to dispell. How will the school in Band A who scored highest in the state in math know that, indeed, it is the best of the group, regardless of parent background? And, how will the school scoring at the 5th percentile in Band E know

whether it is simply one of the lowest among a high scoring group or is one of the lowest among all schools? When the state publishes percentile ranks the incomparability of the comparison bands is readily recognized. When the state does not publish percentile ranks, the band percentile ranks will be interpreted in relative terms: high ranks in low Bands (A-B) will be seen as being lower than low ranks in high Bands (D-E). The actual ranking is unknowable. With the overlap of values of each indicator comparison band, interpretations are misleading and invalid.

What Must Be Done

A valid indicator assesses what it purports to measure and leads to accurate decisions. What do the quality indicators measure? Neither the SB 813 reforms nor state definitions of excellence exist against which the accountability reports may be interpreted. The indicators are confusing, nearly uninterpretable, and lack intrinsic validity. They define education as a product rather than a process and ignore the complexity inherent in teaching human beings. Until the High School Accountability Program publishes data related to worthwhile educational goals based upon a consensual definition of school excellence, it will continue to be a hastily contrived, ill-conceived attempt at evaluating California's public schools.